

JOINT TRACKING OF MOVING OBJECTS WITH EO AND IR CAMERAS

Jing Wang, Hong Man

Department of Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, New Jersey 07030
jwang, hman@stevens.edu

Sachi Desai

US Army RDECOM,
Picatinny Arsenal,
NJ 07806

ABSTRACT

Tracking moving objects, especially human objects in surveillance systems has attracted considerable research attention. This study proposes a novel joint Electro-Optical (EO) and Infrared (IR) cameras tracking approach by employing particle filter. A centroid-based detection technique is used to discover potentially moving objects and obtain the coordinate data. Once moving targets are detected, both EO and IR features are combined to extract object templates for sampling particles. Statistic information of a blob centered at current particle and likelihood of each pixels in terms of foreground, background and occlusion components are obtained, to determine and update importance of each particle and handle temporary occlusion. Hence, particles which can provide accurate prediction are assigned with higher weights. Simulations have been conducted to validate the proposed method.

1. INTRODUCTION

With growing security concerns, it is crucial to develop powerful surveillance systems to monitor, detect and track potentially malicious events and people. However, traditional visible EO video cameras fail to provide accurate information under certain situations, such as illumination variation or occlusion, given their inherent constraints. Intuitively, complementary information can be obtained by deploying devices with complementary features of video cameras, for example, infrared video cameras, which depend on heat dispersion.

IR images are normally of low resolution, therefore shape is a major feature for image data analysis. However, it can monitor objects regardless of light conditions, since IR images are formed on the basis of heat emissivity, conductivity of material surface, as well as reflection, etc.. On the other hand, EO images have color, texture, and shape features of exploration potential. However, EO cameras rely on the sufficient illumination of environment. Therefore, multi-modality surveillance systems have attracted research attentions as possible solutions to improve detection and tracking performance.

Further, many existing military surveillance and monitoring systems are now equipped with both visible light electro-optical (EO) cameras and infrared (IR) cameras. However most of these cameras are operated separately, with EO cameras mainly producing day-light visions, and IR cameras producing night visions. Intuitively we recognize that EO and IR cameras may provide complementary information if they are used jointly. An EO image mostly represents the intensities of reflected visible lights from certain object in the field of view, while an IR image captures the thermal profile of the object.

This study proposes a new technique of tracking moving objects, even with occlusions under weak illumination conditions through joint processing of EO and IR video sequences using particle filters. The article is organized as follows. Section 2 describes the basic ideas and algorithms of proposed moving object detecting and the tracking system. Section 3 presents simulation profile and results. Finally, this article is summarized in Section 4.

2. OVERVIEW OF PREVIOUS WORK

Kang et.al proposed a joint probability model of EO and IR cameras and used Kalman filter to resolve occlusion problem [1] [2]. Kalman filter is an optimal solution of linear systems with Gaussian noise. However, tracking environments are normally complex and may not fit linear system model. Shaohua et al proposed a particle-filter-based tracking system with an appearance-adaptive model [3].

Khan et.al proposed a template-based particle filter system to track ants [4]. Compared with people, ants are more rotation-invariant, hence learning system should be somehow different. James et.al studied multi-target detection and tracking algorithm by deploying boosted particle filter [5]. It depends on color feature of EO images, which are sensitive to light condition. Pupilli et.al developed a particle filter based algorithm to deal with occlusion of objects [6]. It used a large number of particles in state space to track temporarily lost target.

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 01 DEC 2008		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Joint Tracking Of Moving Objects With EO And IR Cameras				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Department of Electrical and Computer Engineering Stevens Institute of Technology Hoboken, New Jersey 07030				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADM002187. Proceedings of the Army Science Conference (26th) Held in Orlando, Florida on 1-4 December 2008, The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 5	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

3. DETECTION AND TRACKING OF MOVING OBJECTS

In this section, the fundamental ideas and algorithms of detecting, as well as tracking moving objects are presented.

General framework of the centroid-based detection and particle-filter-based tracking system is presented as follows,

- Register EO and IR images by transformation and translation algorithms
- Detect moving objects by centroid-based method
- Templates of moving targets are obtained by combining EO and IR counterparts, if exist, to take advantage of heterogenous monitoring mechanism. Otherwise, 1-modality data is deployed
- Particles are generated on detected moving regions to track further changes
- Moving objects are tracked by deploying particle-filter based technique

3.1. Continuous Detection of Moving Objects

EO and IR images were registered by using piece-wise linear transformation. Since it's hard, if not possible, to physically overlap EO and IR cameras to obtain perfectly aligned images, image registration is an essential step for further analyzing the two-modality (i.e. EO and IR) images. The basic approach is to match landmarks, so that the same objects can be overlapped in both images captured by EO and IR cameras respectively [7].

Difference image between sequential frames is derived for further detection of movement. Thereafter, edge detection algorithm is used to obtain boundaries of potentially changing objects. To highlight the region of great intensity change, i.e. possibly with edges, Laplacian filter suitable for black and white images can be applied to facilitate further detection process. Canny edge detection method, which was known to be optimal edge detection algorithm in terms of precision, was used to detect shape information. Shape information is useful for establishing feature correspondence between EO and IR images.

After the boundaries of potentially moving objects are determined, the coordinate values of centroid of the objects are calculated by intersecting the horizontal and vertical scan results. Once the centroid data of changing regions is derived, particle filter tracking algorithm can be triggered to explore suspicious areas.

3.2. Particle Filter

Particle filter is a Monte Carlo method. It is used for non-linear and non-Gaussian problems, which approximates continuous probability density function by using a great number

of samples, i.e. discrete distribution approximation. There are various types of particle filters, such as the sequential importance sampling (SIS), sampling importance resampling (SIR), auxiliary sampling importance resampling, and regularized particle filter [8]. Ellipses are selected to approximate suspicious regions since our main concern is to track human being, which can be generally described by this shape. Shape is a major feature of IR images of relatively low-resolution. Furthermore, ellipse modelling can be extended to track changes of human body contours to obtain more insights [9].

The basic procedures designed for our surveillance-based problem are described as follows,

1. Define state and measurement equations
2. Spread particles on target regions to approximate continuous probability density function
3. Collect historical data to make prediction of next state
4. Predict status of subsequent stage
5. Evaluate real state by measurement techniques
6. Update weights of each particle based on the accuracy of the prediction, or the distance between prediction and actual measurement
7. If degeneracy problem occurs, resampling technique can be used
8. Goto 4

3.3. State Dynamic Models

The following state equation describes human movement dynamics [10]:

$$x = f(x_{t-1}, m_{t-1}) = [r_t, s_t, \dot{r}_t, \dot{s}_t] = \begin{pmatrix} 1 & 0 & \tau & 0 \\ 0 & 1 & 0 & \tau \\ 0 & 0 & a_r & 0 \\ 0 & 0 & 0 & a_s \end{pmatrix}$$

where m is noise vector, (r_t, s_t) denotes image coordinate.

The measurement equation of object edges are as follows:

$$inputx = cx + r_1 \times \cos(\theta)$$

$$inputy = cy + r_2 \times \sin(\theta)$$

Color measurement equation proposed by Yang et.al [11] is presented by

$$(r_i, g_i, b_i) = \sum_{(x,y) \in R_i} ((r(x,y), g(x,y), b(x,y))/A_i)$$

where (cx, cy) , $(inputx, inputy)$ are image coordinates, θ is angel, r_1 and r_2 are ellipse radius.

Movement prediction is updated on the basis of average shift of previous frames.

3.4. Preprocessing Stage

A limited number of samples are placed around the ellipse to approximate posterior probability density function. When sudden change occurs, more particles are required to predict possible movements. Given the object size, 3 concentric ellipses are covered by the particles, i.e. 1 particle per degree, to circle around an object. Based on the property of particle filter, more samples may provide more precise approximation of the actual probability density function. However, resource consumption and computation are more intense as well. In addition, occlusion problem involves even more undeterministic state space change, therefore, more particles are necessary to predict possible behaviors. The next state is predicted by computing state-space equation. In this case, the positions of the ellipses are estimated accordingly to track moving objects.

3.5. Tracking Process by Particle Filter

Weights are computed to highlight the importance of all the particles. Importance is associated with accuracy of approximating continuous probability distribution by discrete measurement. Initially, equal weights are assigned to each particles. After the actual measurement is obtained, the weights of the particles are updated on the basis of the distance between the probability distribution of the template and the region centered at the current particle, as well as the mixture likelihood of foreground, background and occlusion components. Theoretically, the choice of posterior probability distribution is essential for the accuracy of state estimation.

Gaussian mixture model (GMM) is deployed to present objects' or templates' probability density distribution. The experimental results indicate that GMM is an appropriate candidate to evaluate the accuracy of particles in terms of weight update. The following equation defines GMM with more than one components in \mathbb{R}^n for $n \geq 1$ [12]:

$$p(x|\theta) = \sum_{m=1}^M \alpha_m p(x|\theta_m), \forall x \in \mathbb{R}^n$$

where

$$\sum_{m=1}^M \alpha_m = 1, \text{ for } m = 1, \dots, \text{ with } \alpha_m \geq 0$$

and every component follows normal distribution

$$p(x|\theta_m) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Further, expectation-maximization (EM) algorithm is implemented to estimate image distribution. GMM has been a general method to describe density function of image data, since more than one normal distribution component may present the actual distribution more precisely. Histogram of EO and IR segments can be used as an imagery feature of the template and object of interest.

To use the intensity data more efficiently and improve tracking performance, a mixture appearance model is incorporated into the particle filter [3]. The mixture model consisting of background, template, and occlusion elements can provide a more effective particle sampling, as well as handle occlusion case by expanding the observed area. If a potential occlusion is inferred, more particles are necessary to accurately track further movements of the observed objects. The weight of each pixel is updated based on the mixture likelihood, given the three components. In this way, particles are sampled around the area with higher likelihood of reappearance of the object.

Joint expectation maximization of EO and IR is a promising method to take advantage of multi-modality. Further, RGB color feature of EO images is a good compliment for IR images in terms of visual properties, since resolution of IR images are generally lower. However, IR images are less susceptible to light change, since thermal cameras are susceptible to heat emission, instead of light conditions. In addition, IR images can provide a foundation for simple and accurate change analysis, given the empirical results. Once targets are localized, shape correspondence is the key to relate EO and IR image data. Visual features of EO images can be combined and explored to obtain more insights of observed scenario.

After statistic data of template and particle-centered regions is derived, the distance between probability distributions is computed to determine weights. Kullback-Leibler Divergence (KLD) measures the distance between the original probability distribution and a candidate probability distribution [13]. It is defined as

$$D_{KL}(P, Q) = \sum_i P(i) \log P(i) - P(i) \log Q(i)$$

where P is probability density function of the templates, and Q represents the particle's probability density function.

As to the proposed tracking framework of particle filter, the distance between probability distribution of the template, i.e. the original one, and that of the region centered at a particular particle is measured. Theoretically, particles close to the true centroid have similar probability distributions, and therefore deserve higher weights to effectively allocate limited resources.

Weights associated with each particles are calculated by

$$w_t^i = w_t^i p(y_t | x_t^i)$$

Thereafter, weights are normalized as

$$w_t^i = w_t^i / \gamma, \text{ where } \gamma = \sum_{i=1}^N w_t^{(i)}$$

Posterior probability distribution is calculated by [10]

$$p(x_t | z_{1:t}) = \sum_{i=1}^N N_s w_t^i \delta(x_t - x_t^i)$$

where x_t is state at time t ; $z_{1:t}$ denotes history of observation; and w_t^i are the weights of particles.

In the case of sample degeneracy, resampling technique is deployed [10].

4. TESTING DETECTION AND PARTICLE-FILTER-BASED TRACKING SYSTEM

In the previous section, the detection and tracking techniques of moving targets are discussed. Simulation procedure, results, and implications are presented in this section.

Detection system and particle-filter-based tracking system were implemented and tested by Matlab. OTCBVS04 DATASETS [14] were selected to conduct simulation, since both EO and corresponding IR images are provided for further exploration. Temporary occlusion case was tested by playing back old frames and remixed frames, while weak illumination region was selected for testing. Testing procedure followed the proposed detecting and tracking technique. Simulation results in Figure 1, 3, and 4 show that moving parts of human body can be detected and tracked. However, there was a false alarm on detecting phase, where noise was captured.

Fig. 1. Tracking Moving Parts of Human Body



Particles making greater contribution to predictions gain more weights than others, hence limited resources can be used more efficiently. Further, changing parts of body, such as legs attract more attention. Temporary occlusion under weak illumination case was handled by placing more particles. Likelihood function with the three components is helpful in determining weights of particles. State dynamics considering previous movements are more efficient in prediction. The testing result is shown in Figure 2. Hence, it can be used to obtain more details on moving objects. This work will be extended by deploying more ellipses to describe contour of human body and monitoring interacting targets in future.

5. CONCLUSION

In conclusion, this work introduced a new multi-modal object tracking method based on particle filters. A movement detec-

Fig. 2. Tracking Output after Temporary Occlusion and the Statistic of the Target

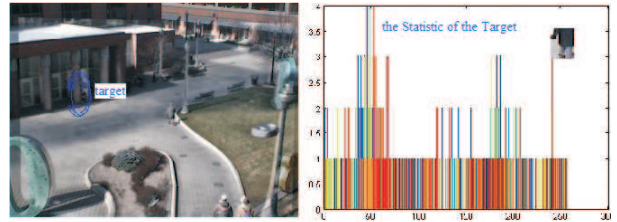


Fig. 3. Tracking Results of the First Sequence

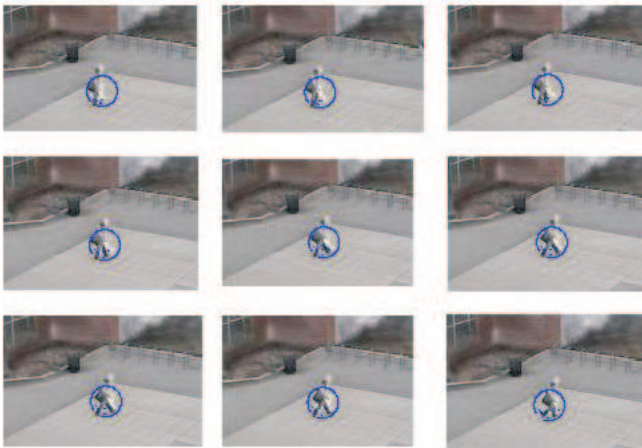


tion and tracking system based on joint EO and IR cameras was developed. Centroid-based algorithm was implemented to detect the changes and collect the templates. Particle filter was used to track moving objects, which aimed at resolving temporary occlusion with insufficient illumination problem. Our study has shown that particle filter appears to be a promising mathematical framework for multi-modal data fusion, in which observations and features from different modalities are used to estimate the joint posterior probability at each tracking stage. Simulation results showed that the system can detect moving object, and track it, even under temporary occlusion case. This approach can be more efficient than the simple integration of separate tracking results from individual modalities, and it may have great potential in most military multi-sensor systems. One of our future works is to extend this method for acoustic and visual sensor data fusion.

6. REFERENCES

- [1] J. Kang, I. Cohen, and G. Medioni, “persistent objects tracking across multiple non overlapping cameras,” in *WACV-MOTION '05: Proceedings of the IEEE Workshop on Motion and Video Computing (WACV/MOTION'05) - Volume 2*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 112–119.
- [2] J. Kang, K. Gajera, I. Cohen, and G. Medioni, “detection and tracking of moving objects from overlapping

Fig. 4. Tracking Results of the Second Sequence



EO and IR sensors”, in *“Object Tracking and Classification Beyond the Visible Spectrum”*, 2004, p. 123.

- [3] S. K. Zhou, R. Chellappa, and B. Moghaddam, “Visual tracking and recognition using appearance-adaptive models in particle filters.” *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1491–1506, 2004.
- [4] Z. Khan, T. Balch, and F. Dellaert, “An mcmc-based particle filter for tracking multiple interacting targets,” 2004.
- [5] K. Okuma, A. Taleghani, N. de Freitas, J. J. Little, and D. G. Lowe, “a boosted particle filter: Multitarget detection and tracking”, in *European Conference on Computer Vision*, 2004, pp. Vol I: 28–39.
- [6] M. Pupilli and A. Calway”, “real-time camera tracking using a particle filter”, in *Proceedings of the British Machine Vision Conference*. BMVA Press, September 2005, pp. 519–528.
- [7] H. E. A. . R. M. J. T. ”Lei, B. J., “reviewing camera calibration and image registration techniques”, Telematica Instituut, Tech. Rep. TI/RS/1999/113, 1999.
- [8] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp”, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” *Signal Processing, IEEE Transactions on*, vol. 50, pp. 174–188, Feb. 2002.
- [9] M. Iiyama”, “pose estimation of human body on low resolution voxel data”, Kyoto University, 1997.
- [10] Y. C. ”Yong Rui, “better proposal distributions: Object tracking using unscented particle filter”, in *CVPR (2)*, 2001, pp. 786–793.
- [11] R. D. ”Changjiang Yang and L. Davis”, “fast multiple object tracking via a hierarchical particle filter”, in *“ICCV ’05: Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1”*. Washington, DC, USA: IEEE Computer Society, 2005, pp. 212–219.
- [12] Y. W. X. Yang; and K. L. Chan, “Unsupervised color image segmentation based on Gaussian mixture model,” *Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, vol. 1, no. 15-18, pp. 541 – 544, Dec 2003.
- [13] [Http://www.wikipeida.com](http://www.wikipeida.com).
- [14] J. I. W. on Object Tracking and C. B. the Visible Spectrum (OTCBVS 2004)”. Washington, D.C: OTCBVS04, June 2004.